# Architectural Issues
# or
# Making Sense of the Zoo

http://www.slac.stanford.edu/~abh/PPDG/Zoo.html

**Andrew Hanushevsky**

**Stanford Linear Accelerator Center**

Produced under contract DE-AC03-76SF00515 between Stanford University and the Department of Energy



## Architectural Issues

- Replication
  - How do we provide for a multi-cultural model?
    - Solves the immediate problem
    - Encourages creative solutions
- Security
  - How do we provide for a low-cost security model?
    - Solves the immediate problem
    - Doesn't eat us administratively alive
- Replica Catalog
  - How do we provide for a scalable model?
    - Solves the immediate problem
    - Won't fall apart once beyond tinker-toy use

## Replication Issues

- There are (at least) *two* distinct replication contexts
  - Wide Area Replication (WAR)
    - Replication of files between “sites” (e.g., SLAC, IN2P3, etc)
  - Local Area Replication (LAR)
    - Replication of files within a “site”
- Each context has it’s own peculiar requirements
  - Leads to different approaches on replication management



## WAR vs LAR

- Primary reason for replication differs
  - WAR tries to duplicate data at geographically remote sites
    - Availability driven
      - Client-directed performance criteria
  - LAR tries to duplicate data among local hosts
    - Performance driven (e.g., dynamic load balancing)
      - Server-directed performance criteria
- Frequency differs
  - WAR is typically less frequent than LAR
    - Though when it happens it happens en-masse
- Network reliability and speed differs
  - WAR networks are less reliable, slower and have higher latency



## One Size Fits All?

- One size fits all solutions are problematic
  - WAR-oriented replication is generally heavy-weight
    - Availability is the most important issue
    - Deliberate contractual replication decisions
  - LAR-oriented replication is generally light-weight
    - Performance is the most important issue
    - Instantaneous automatic replication decisions
- One size fits all solution should not be forced
  - Indeed, our direction gravitates towards multiple solutions
- How can this be easily accomplished?
  - Want the zoo of solutions to be admired rather than abhorred

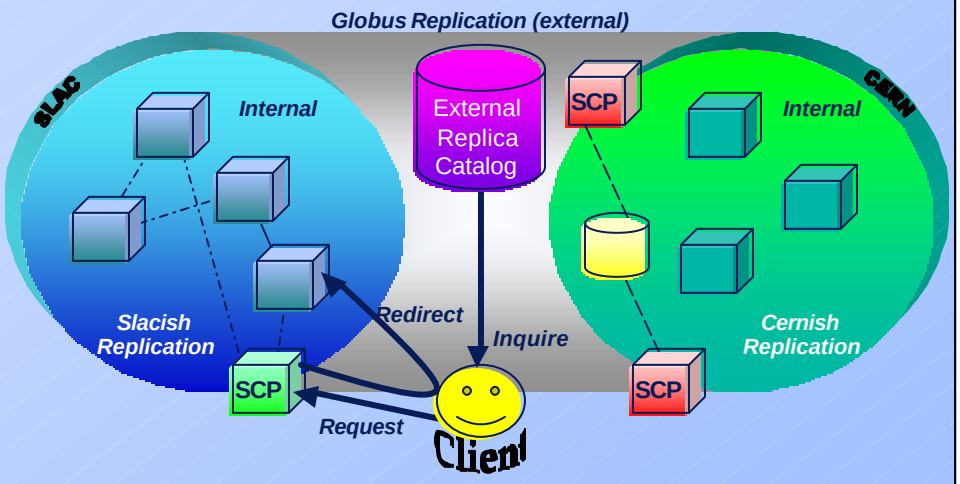


## An Architectural Proposal

- Differentiate the notion of
  - Inter-site or external replication, and
  - Intra-site or internal replication
- A site is an “arbitrary” collection of machines
- External Replication
  - Replicas tracked to a site
    - One or more boundary hosts or site contact points (scp)
- Internal Replication
  - Replicas tracked to a particular host within a site
    - The boundary host or scp provides in-site navigation support
- In short – Autonomous Replication



## Autonomous Replication






## Autonomous Replication Advantages

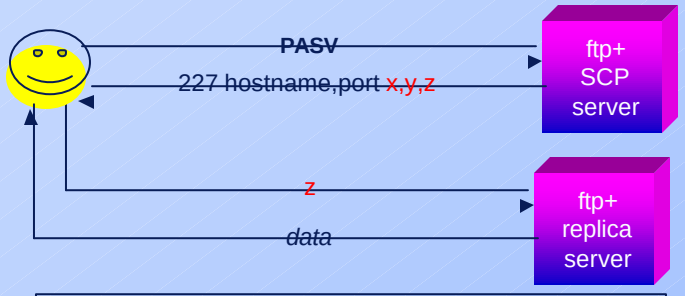
- Natural peer-to-peer architecture
  - Each site is independent but can cooperate as needed
- Does not limit replication technology R&D
  - Each site can research and deploy site-appropriate strategies
    - Overall replication environment is not impacted
    - Naturally explains the various replication strategies
- Compatible with Globus and SRB technology
  - Makes use of the current protocol redirection capabilities
    - GSI-ftp+
    - http
  - External replication may be cascaded into internal replication
    - You can use any technology that supports ftp or http

## Autonomous Replication Implementation

- External replication via Globus API's
  - Can continue with current track
- Internal replication via site-specific mechanism
  - Can be Globus or any other SCP-compatible mechanism
- SCP bridges the two worlds in one of two modes
  - Compatibility Mode
    - Performs expected functions of standard ftp/http server
  - Extended Mode
    - Implements complete redirection protocol
  - Can use both modes on a request-specific basis
  - Fully compatible with Globus and SRB



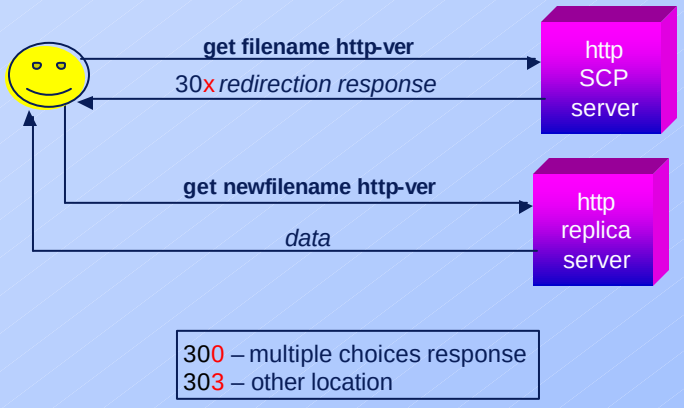
## SCP ftp+ Compatible Redirection Protocol




x – optimal tcp buffer size
y – optimal number of data streams
z – scp-specific information to be sent on data connection

*not caste in concrete*

## SCP http Redirection Protocol




300 – multiple choices response
303 – other location



## Security Architectural Issues

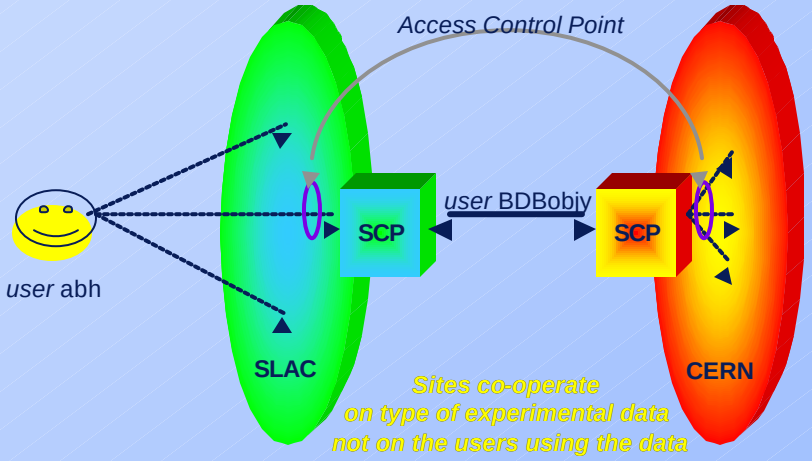
- Current replication system (I.e., Globus) relies on PKI
  - Difficult to administer and very labor-intensive
  - Yet another security infrastructure to deploy and maintain
- Changing the security model is difficult
  - Politically
    - No agreement on the best security model (e.g., Kerberos?)
  - Technically
    - Requires major extensions to existing systems (e.g., Globus)
- The “best” solution is to change the processing model
  - This is a management issue with technical implications

## The Service Model

- Provide a data service to multiple users via agents
  - Users never directly access data outside their site
    - Need installation-specific authentication within the site
    - Access to data outside the site is via a named service agent
    - Remote access control based on the agent name
      - No need to support delegation
  - Very small number of well identified agents
    - Small number of certificates to manage
    - One agent for a particular type of managed data
      - BaBar Objectivity databases
- This is *not* a general solution to data access
  - PPDG does not need a general solution
    - We have a well constrained data access problem
- It greatly simplifies security without undermining it

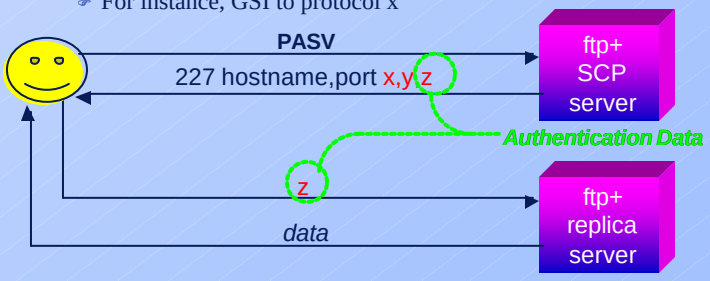


## Security in the Service Model



## Further Lightening Security via Transforms

- Service model solved many problems but not all
  - Still need every data server to be a PKI heavy-weight
- SCP redirection protocol allows for security transforms
  - A transform is a substitution of one security model for another
  - Server directed at destination site
  - The ftp+ and http redirection models provide for transforms
    - For instance, GSI to protocol x






## Replica Catalog Architectural Issues

- Need a robust scalable catalog
  - Many LDAP implementations are not scalable (e.g., Open LDAP)
  - Commercial LDAP servers too expensive (e.g., Oracle at $500K+)
- Solutions are not easy
  - Need to identify minimum set of information to place in catalog
    - Prevent catalog bloat, the largest impediment to scalability
  - Develop an SQL LDAP back-end?
    - Compatible with Oracle and other database vendors.
  - Develop an Objectivity LDAP back-end?
  - Spend the big bucks
    - Still need objective evaluations on available products

## Conclusions

- Autonomous Replication
  - Provides for diverse systems without requiring them
  - Fully compatible with Globus and SRB
  - Captures the HEP R&D model
    - Not necessarily bad
- Service Security Model
  - Eases the administrative overhead of PKI
  - Adequate for most HEP endeavors
  - Allows for protocol transforms
    - Easy to maintain site-specific security
- Replica Catalog
  - No solutions in site, sorry to say